

Evolúció és rendszerbiológia

MTA Doktori Tézis

Pál Csaba

Szintetikus és Rendszerbiológiai Egység

Magyar Tudományos Akadémia Szegedi Biológiai Kutatóközpont

Szeged, 2017

A XX. század elején a evolúcióbiológiát a mendeli genetikával ötvöző modern evolúciós szintézis számos biológiai jelenség megértését tette lehetővé. Egyesített ezen kívül olyan korábban elhatárolt tudományterületeket, mint a biogeográfia, taxonómia, ökológia és populációgenetika. Hatalmas sikere ellenére azonban bizonyos kulcsfontosságú kérdések továbbra is megválaszolatlanok maradtak. Ezek tisztázásához új megközelítést jelent a rendszerbiológia¹. E tudományterület célja az evolúcióbiológia központi problémáinak felülvizsgálása, amelynek érdekében a molekuláris hálózatok részletes vizsgálatát, valamint sejtes alrendszerek számítógépes modellezését ötvözi a populációgenetika módszereivel és a kísérleti evolúcióval¹.

A rendszerbiológia és az evolúció közötti elvi és módszertani kapcsolatokat a továbbiakban négy fő témakörre koncentrálni mutatom be. Az első a génkiütés paradoxonhoz kapcsolódik. Miért tűnik nélkülözhetőnek a legtöbb gén? Ezen kérdés kifejtése vezet a második témakörhöz, a kompenzációs evolúció problémájához. Ezt követően bemutatom a bakteriális genommérnökség legújabb fejlesztéseit, és azt, hogy ezek hogyan használhatók fel a bakteriális genomok evolúciójának laboratóriumi vizsgálatára. Végezetül, röviden összefoglalom a mikrobákban megfigyelhető antibiotikum-rezisztencia és kollaterális szenzitivitás témakörében végzett kutatásainkat.

I. A génvesztés evolúciója

A legtöbb organizmus számára egyetlen gén elvesztése nincs hatással a túlélésre². Az élesztőben (*Saccharomyces cerevisiae*) az egyszeres géndelációk hatása alapján, a gének mindössze 20%-a tűnik esszenciális, de hasonló megfigyelésekről számolnak be a fonálféreg, *Caenorhabditis elegans*, a *Bacillus subtilis*, és számos más faj esetén is (1. táblázat).

organizmus	fehérjekódoló gének száma	esszenciális gének becsült %-os értéke
<i>M. genitalium</i>	482	79
<i>H. pylori</i>	1590	17
<i>H. influenzae</i> Rd	1850	38
<i>S. aureus</i> N315	2594	25
<i>M. tuberculosis</i> H37Rv	4402	15
<i>B. subtilis</i>	4099	6.6
<i>E. coli</i> K12	4296	7.1
<i>S. typhimurium</i> LT2	4597	11
<i>C. glutamicum</i> R	3052	22.6
<i>F. novicida</i> U112	1719	23
<i>S. cerevisiae</i>	5794	19
<i>C. elegans</i>	19 099	7

1. táblázat Esszenciális gének eloszlása különböző modellorganizmusok között. A táblázat forrása: ².

Ezek a megfigyelések számos kérdést vetnek fel: A gének nagy része valóban nélkülözhető az organizmus számára? Hogyan képes fenntartani a növekedését egy géndelációs törzs?

Ezekre a kérdésekre kerestük a választ 2004-ben³ az élesztő (*Saccharomyces cerevisiae*) genomszintű metabolikus hálózatának modellezése segítségével. Korábbi biokémiai tanulmányok alapján rekonstruáltuk a 809 anyagcsereterméket és 851 ezeket összekötő biokémiai reakciót tartalmazó hálózatot³. Ezt követően oly módon optimalizáltuk az anyagcsere-hálózatot, hogy a 282 előre meghatározott, ökológiailag releváns tápközegben előállítsa a

növekedéshez szükséges anyagcseretermékeket. A modell alapján a legtöbb gén esetén elmondható, hogy annak elvesztése a 282 vizsgált körülmény csak igen kis hányadában (10%) jár súlyos költséggel. Ebből tehát az következik, hogy a legtöbb gén megléte csak bizonyos környezetben nélkülözhetetlen³.

Ezt a felvetést több kísérletes tanulmány is alátámasztja. Élesztőben a biokémiai reakciók fluxusának közvetlen mérésével kimutatták, hogy a látszólagosan nélkülözhető gének kb. 50 százaléka inaktív laboratóriumi körülmények között^{4,5}. Ezenkívül egy nemrégiben megjelent nagy áteresztőképességű kemogenomikai tanulmány szerint az élesztő 5000, látszólag nélkülözhető génjének 97%-a legalább egy környezetben hozzájárul a túléléshez⁶. Jellemző továbbá, hogy a káros fenotípusok általában a vizsgált környezetek kis hányadára korlátozódnak⁶.

Az említett megfigyelések nem zárják ki a gének nélkülözhetősége mögötti egyéb mechanizmusokat⁷. Egy gén elvesztése kétféle módon kompenzálódhat: a gén redundáns funkciójú változata által, vagy az anyagcsereútvonalak átrendezésével, alternatív útvonalak révén^{3,7}. Annak tisztázására, hogy a két mechanizmus közül (génduplikáció vagy alternatív útvonalak) melyik a gyakoribb, újra az élesztő metabolizmusának vizsgálatához fordultunk³. Azokra az esszenciális enzimikus reakciókra fókuszáltunk, amelyek deléciója a növekedés gátlásához vezet. Ez alapján úgy becsüljük, hogy a génduplikációk 15-28%-ban, míg az alternatív metabolikus útvonalak csak 4-17%-ban felelnek a gének nélkülözhetőségéért. Később ugyanebben a fajban elvégezték a biokémiai reakciók fluxusának mérését, amely azonos eredménnyel szolgált. A kísérlet alapján a 207 életképes törzs esetében, amely aktív reakciókat érintő mutációkat hordoz, a redundáns funkciójú duplikálódott gének jelentik a nagyobb (75%), és az alternatív útvonalak a kisebb (25%) hányadát a hálózatok robusztussága mögötti molekuláris mechanizmusoknak. Ezek az eredmények azonban nem zárják ki, hogy más sejtes rendszerekben a metabolikus útvonalak átrendezése a gyakoribb.

Ezt követően feltettük a kérdést, hogy a génduplikáció (vagy alternatív útvonal) vajon azért terjedt-e el és maradt fent, mert más mutációk káros hatásait kompenzálja³. Korábbi elméleti kutatások szerint a természetben megfigyelhető mutációs ráta és populációméret mellett, a legtöbb organizmusban nem alakulhatnak ki a mutációk káros hatását kivédő változások. Miért fordulnak elő mégis duplikációk a genomokban melyek átfedő biokémiai reakciókat kódolnak (izoenzimek)? Az élesztő anyagcsere-hálózatának elemzése kimutatta, hogy az izoenzimek nem katalizálnak gyakrabban esszenciális reakciókat, mint nem esszenciálisakat. Ehelyett, jelenlétük a metabolikus hálózat azon részein jellemző, amelyeken magasabb fluxusra van szükség. Ebből az következik, hogy a duplikációk fennmaradásának oka, hogy növelik a fluxust, amellyel szelektív előnyt biztosítanak, valamint a fluxus növekedésének másodlagos következményeként kompenzálják a génvesztést.

A hálózatok robusztusságának kérdését úgy is megközelíthetjük, hogy olyan gének evolúcióját vizsgáljuk, amelyek szekvenciájuk alapján nem mutatnak rokonságot, de képesek kompenzálni egymás nullmutációit. Az ilyen genetikai interakciók minimum 51 százalékáról elmondható, hogy az egymást kompenzáló mindkét gén deléciója csak bizonyos környezetben letális⁸. Ezek az eredmények egyeznek azzal a modellel, amely szerint az enzimek bizonyos tápanyag jelenléte és nem a mutációk hatásának kivédése miatt esszenciálisak.

II. Kompenzáló evolúció

Az emberi populációkban meglepően nagy számban fordulnak elő genetikai eredetű megbetegedések⁹. Gyakori azonban, hogy ugyanaz a káros mutáció különböző egyéneknél más-más tünetekkel jár, vagy egyáltalán nincs hatása. Jellemző továbbá, hogy az emberben káros hatással rendelkező mutációk konzerváltak más, közeli rokon fajokban^{10,11}. Mi lehet ennek az oka? Egyik lehetőség az, hogy az evolúciós alkalmazkodással együtt jár olyan mutációk felhalmozódása, amelyek

káros pleiotróp hatással rendelkeznek. Így tehát az evolúció során az organizmusokban bekövetkező jelentős változások nem csak az új környezethez történő alkalmazkodást, hanem az adaptív mutációk káros hatásának kivédését is szolgálják. Az 1920-as évek elején Ronald Fisher leírta azt az elméletet, miszerint az adaptáció hasznos mutációk fokozatos felhalmozása révén történik¹². Ezzel szemben, Sewall Wright azt a nézetet vallotta, hogy a bizonyos feltételek mellett káros mutációk mérföldkőként szolgálnak az adaptáció során, mivel olyan evolúciós útvonalak elérését teszik lehetővé, amelyek más esetben nem elérhetőek¹³. Az elmúlt évtizedekben tömegével születtek elméleti munkák a témában, azonban kísérletes adatok – különösen genom szinten – nehezen fellelhetőek.

Munkánk során¹⁴ ennek a problémának egy kevésbé kutatott aspektusára fókuszáltunk. Feltettük a kérdést, hogy a káros következménnyel járó génvesztés elősegíti-e az adaptációt, illetve mik lehetnek ennek a folyamatnak a káros mellék következményei¹⁴. Célunk elérése érdekében többféle módszertani megközelítést alkalmaztunk: laboratóriumi evolúciós kísérletet és genomikai analízist kombináltunk bioinformatikai módszerekkel és részletes molekuláris elemzéssel¹⁴. A vizsgálatok során számos újszerű következtetésre jutottunk:

A génvesztést követő kompenzáló evolúció általános. A káros, de életképességet nem befolyásoló deléciók legalább 68%-a a genom más részein bekövetkező adaptív mutációk révén kompenzálható.

Az elveszett molekuláris funkció teljes helyreállítása ritka. A teljeskörű genomikai és fenotípusos analízis során kimutattuk, hogy az evolválódott vonalak eltérnek egymástól és új adaptív csúcsokat értek el. A vad típus fiziológiás állapota tehát általában nem áll helyre, valamint gyakoriak a káros pleiotróp hatások.

Eredményeink alapján úgy véljük, hogy a fajok közti génkészletbeli különbségek a részben a kompenzáló evolúció hatásai és nem feltétlenül a környezethez való alkalmazkodást tükrözik.

III. Evolúciós genommérnökség

A teljes genomra kiterjedő genetikai mérnökség lehetőséget ad specifikus genomi régiók irányított és kombinatorikus szerkesztésére¹⁵. Az e területen elért legfrissebb újítások példátlan lehetőséget nyújtanak előre meghatározott funkcióval rendelkező, komplex molekuláris útvonalak tervezéséhez¹⁶. Az ebben a témában született eddigi tanulmányok egy része olyan új útvonalak tervezésére helyezte a hangsúlyt, amelyeken keresztül gyógyászati és ipari célból hasznosítható molekulák állíthatók elő, míg mások további tervezésre alkalmas genomi konstrukciókat hoztak létre.

Azonban a genommérnökség az evolúció valós idejű tanulmányozásához is lehetőséget biztosít, kiküszöbölve a jelenleg használt evolúciós vizsgálati módszerek limitációit¹⁸. A módszer alkalmas többek között a) kiterjedt genomi régiók, vagy teljes kromoszómák gyors szerkesztésére és irányított evolúciójára, b) rövid DNS szegmensek (promóterek, kódoló régiók) vagy teljes genomok szintézisére és kombinatorikus összeillesztésére, c) kiterjedt régiók vagy akár teljes genomok kémiai szintézisére és új gazdaszervezetbe illesztésére¹⁷.

A genommérnökség jelenleg három fő limitációval rendelkezik: i) csak néhány laboratóriumi modellszervezetben alkalmazható (mint pl. az *Escherichia coli*), ii) a gazda genomjának jelentős mértékű, előzetes módosítását igényli, valamint iii) számos nem kívánt („off-target”) módosítás felhalmozásához vezethet, amelyek száma gyakran felülmúlja a kívánt mutációk számát. Az említett problémák nagy mértékben megnehezítik a technika széleskörű biotechnológiai alkalmazását.

Legfrissebb kutatásaink során ezekre kívántunk megoldást találni^{19,20}. A multiplex automatizált genommérnökség (MAGE) korábbi fejlesztéseit alapul véve²¹, kifejlesztettünk egy egyszerű és kompakt megoldást. Ehhez első lépésként jellemeztünk egy olyan domináns mutációt, amely a baktériumokban található metiláció közvetített mismatch hibajavítás (methyl-directed mismatch repair=MMR) egyik kulcsfehérjéjét érinti. Ennek segítségével revezibilisen inaktíválható a hibajavítás a célsejtekben²⁰. Ezt kihasználva kifejlesztettünk egy új genommérnöki eljárást, amelyet pORTMAGE-nek neveztünk el. A módszerről bebizonyítottuk, hogy alkalmas nagy áteresztőképességű genomszerkesztésre, egyszerre több lókuszt hatékonyan módosítása révén.

A pORTMAGE eljárással módosított törzsekben a teljes genomszekvenálás során nem kívánt, off-target mutációt nem azonosítottunk²⁰. Ez komoly előnyt jelent a korábbi módszerekhez képest. Ezenkívül rendszerünkkel megoldottuk a hordozhatóság problémáját is. Mivel az MMR rendszer konzervált, a domináns mutáció átvihető különböző baktériumfajokba. Így, a teljes genomszerkesztést szolgáló szintetikus operont egy széles gazdaspecificitású vektora helyezve, a módszer számos fajban használhatóvá vált.

Munkánk során a pORTMAGE rendszert sikeresen alkalmaztuk biotechnológiailag és klinikailag jelentős enterobaktériumokban²⁰. A vizsgálatban antibiotikumokkal szemben rezisztenciát okozó mutációk hatását tanulmányoztuk *Salmonella entericában* és *Escherichia coliban*. Kimutattuk, hogy annak ellenére, hogy a két faj kb. 100 millió évvel ezelőtt vált el egymástól, a mutációk hatása általános konzervációt mutat. Ez az eredmény jövőbeli szisztematikus tanulmányok alapját képezi²⁰.

Összefoglalva, a rendszerünk alkalmas arra, hogy mindössze egyetlen transzformációval enterobaktérium fajok széles skáláját tegye alkalmassá a genomszerkesztésre, kiküszöbölve a nem kívánt mutációk felhalmozását.

Fejlesztésünk széleskörűen alkalmazható, többek között bioszintetikus útvonalak gyors optimalizálására, és ezen keresztül hasznos molekulák előállítására számos baktériumfajban. Segítségével kikerülhetővé válik a fáradságos laboratóriumi optimalizálás folyamata. Ezen felül, kísérleteink alapján úgy véljük, hogy munkánk új utat nyithat meg különféle kutatási területeken, mint például a funkcionális genomika és az evolúcióbiológia. Módszerünk, a pORTMAGE, elsőként alkalmas a mutációk hatásának szisztematikus összehasonlítására és az episztázis vizsgálatára különböző baktériumfajokban.

IV. Antibiotikum-rezisztencia és kollaterális szenzitivitás

Az alap és alkalmazott kutatás számára egyaránt kiemelkedő jelentőséggel bír annak megértése, hogy a mikrobák rezisztenciájának kialakulása egy adott antibiotikummal szemben hogyan növeli (keresztrezisztencia) vagy csökkenti (kollaterális szenzitivitás) az életképességet más szerekkel szemben²². Nyilvánvaló klinikai jelentősége ellenére, nagyléptékű tanulmányok hiányában, csak korlátozott ismeretanyaggal rendelkezünk erről a kérdésről.

Kutatásunk során *Escherichia coliban* feltérképeztük a különböző antibiotikumok közötti keresztrezisztencia és kollaterális szenzitivitás kapcsolatokat, és felderítettük az ezek mögött meghúzódó molekuláris mechanizmusokat. A munka során laboratóriumi evolúciós kísérletet kombináltunk genomszekvenálással és funkcionális analízissel. Az evolúciós kísérlet során az *E. coli* K12 törzs populációit párhuzamosan adaptáltattuk egy-egy (összesen 12-féle) klinikailag releváns antibiotikummal szemben, fokozatosan emelve az antibiotikum koncentrációját. Ennek eredményeképpen a vad típushoz képest maximálisan 300-szoros növekedést értünk el a minimális gátló koncentrációban (MIC). Az elért rezisztencia mértéke minden esetben elérte vagy meghaladta az EUCAST által meghatározott klinikai határértékeket, valamint az evolált törzsek 52%-a egyszerre több antibiotikummal szemben is rezisztenciát mutatott. Következő lépésként teszteltük az evolált populációk érzékenységét a többi 11 antibiotikummal szemben, amelynek eredményeként feltérképeztük az antibiotikumok közötti keresztrezisztencia-kölcsönhatásokat. A továbbiakban, az ezek mögött meghúzódó molekuláris mechanizmusok felderítése érdekében az evolált vonalakon teljes genomszekvenálást és biokémiai tesztek végeztünk.

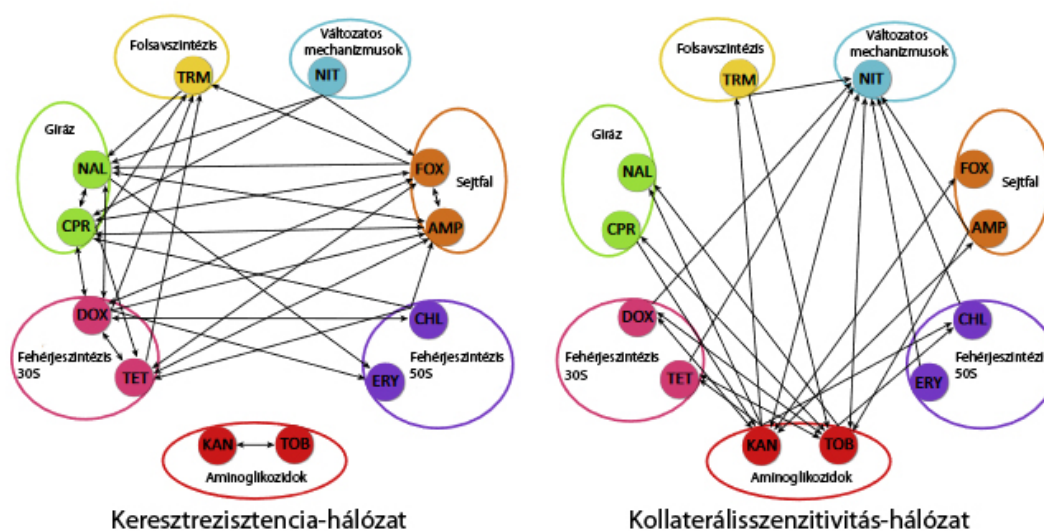
A vizsgálatokból az alábbi következtetéseket vontuk le:

A) A keresztrezisztencia-hálózat sűrű, amely arra utal, hogy egy antibiotikumhoz való alkalmazkodás gyakran okoz multirezisztenciát.

B) A hálózatban gyakoriak az aszimmetrikus kapcsolatok: az „A” stresszforrással

szembeni védelem „B” stresszforrással szemben is véd, de ugyanez fordítva nem igaz.

C) A keresztrezisztencia-hálózat az antibiotikumok tulajdonságai alapján nagy mértékben előre jelezhető.



1. ábra A laboratóriumban evolált baktériumok antibiotikumokkal szembeni érzékenységének meghatározása alapján két hálózat rajzolható fel. Az „A” antibiotikumtól a „B”-be mutató nyíl jelzi, hogy az „A” elleni rezisztencia növelte (kollaterálistenzitivitás) vagy csökkentette (keresztrezisztencia) a „B”-vel szembeni érzékenységet. Forrás: Pál és kollégái 2015.

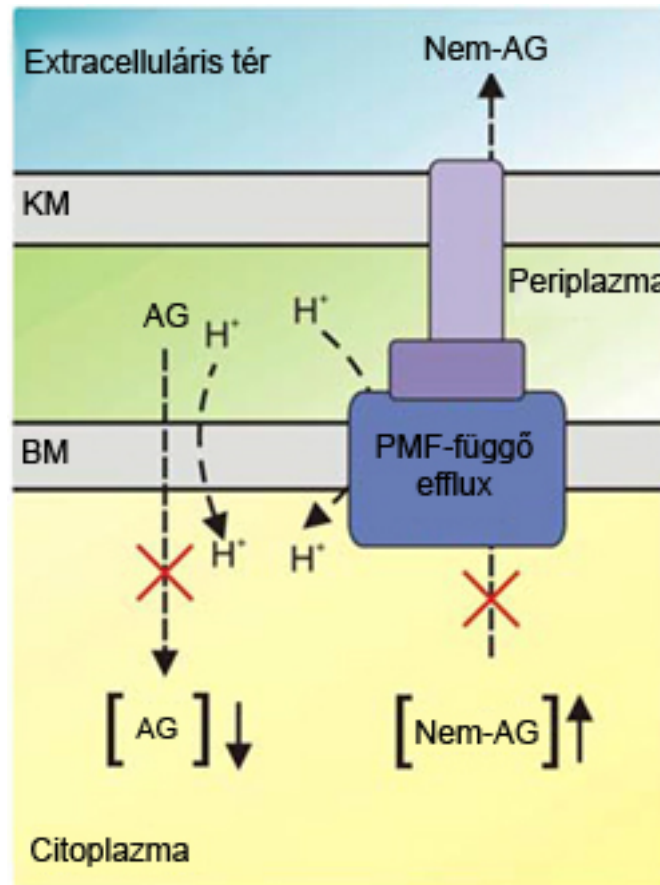
Kísérleteink felfedték a molekuláris szinten zajló párhuzamos evolúció nyomait. A különböző antibiotikumokhoz alkalmazkodott törzsekben párhuzamosan azonos mutációk jelentek meg, amelyek több szerrel szemben is rezisztenciát biztosítottak²³. A keresztrezisztenciáért felelős mutációk sokféle molekuláris mechanizmust érintettek. Az érintett gének között találtunk multidrog efflux pumpát kódoló és metabolikus géneket, valamint a) oxidatív, b) tápanyag- és c) membránstressz elleni

védekezésben szerepet játszó géneket is. Ezen felül arra is fény derült, hogy a keresztrezisztencia-mintázat kialakításához nagy mértékben hozzájárulnak a globális transzkripciót szabályozó gének, amelyek a transzkripciós mintázat genomszintű átrendezéséért felelősek.

Eredményeink közül talán a leginkább figyelemre méltó az a felismerés, hogy a keresztrezisztenciáért felelős mutációk széleskörű pleiotróp hatással rendelkeznek^{23,24}. A különféle faktorokkal szembeni egyidejű védettség tehát általános lehet²⁵. Így a természetben előforduló egyéb stresszfaktorokhoz történő alkalmazkodás melléktermékeként növekedhet az antibiotikumokkal szembeni tolerancia is.

Kollaterális szenzitivitás és multirezisztencia a baktériumokban

Korábbi tanulmányok kimutatták, hogy az egy antibiotikummal szemben kialakuló rezisztencia gyakran jár együtt számos más szerrel szembeni növekedett ellenállóképességgel. Keveset tudunk azonban a kollaterális szenzitivitás előfordulásának gyakoriságáról. Kollaterális szenzitivitásról akkor beszélünk, ha a rezisztencia kialakulása egy másik szerrel szembeni növekedett érzékenységgel párosul. Kutatásaink kimutatták, hogy ez a jelenség igen gyakori²⁴. Az e mögött meghúzódó mechanizmusok megértése azonban további kutatást igényel. Egy lehetséges magyarázat lehet a sejtek membránpotenciáljának megváltozása, amely bizonyos antibiotikumokkal szemben rezisztenciát, míg másokkal szemben túlérzékenységet okoz. Ezt laboratóriumi és klinikai tanulmányok egyaránt leírták²⁴. Ezek alapján ellentétes mechanizmusok létét feltételezhetjük, amelyek az antibiotikumok intracelluláris koncentrációját a membránpolaritás megváltoztatása révén módosíthatják.



2. ábra Kollaterális szenzitivitást okozó mechanizmus. A membránpotenciál belső membránon keresztüli megváltozása kettős hatással rendelkezik: csökkenti számos aminoglikozid-típusú antibiotikum felvételét, valamint mérsékli a PMF-függő efflux pumpák aktivitását. Forrás: Lázár és kollégái 2013.

Klinikai jelentőség

A keresztrezisztencia/kollaterális szenzitivitás hálózat térképe forrásként szolgálhat a gyógyászat számára és elősegítheti a megfelelő terápia kiválasztását. Az optimális antibiotikum-kombináció kiválasztása például nagyban függhet a szerek közötti fiziológiai kölcsönhatástól és azoktól a mutációktól, amelyek mindkét szerrel szemben ellenállóképességet biztosítanak. Kimutatták, hogy két antibiotikum közötti keresztrezisztencia csak nagyon kis mértékben függ attól, hogy kombinációban alkalmazva fennáll-e közöttük szinergizmus²⁶. Az antibiotikumok közti

keresztrezisztenciáról és szinergizmusról gyűjtött minél nagyobb mennyiségű további információ különösen hasznos lehet a jövőben a multirezisztens kórokozók elleni terápia kialakításában. Sokat vitatott kérdés, hogy egy adott antibiotikum lecserélése, majd újra felhasználása okozhatja-e a rezisztencia megjelenését²⁷. Munkánk eredményei azt jelzik, hogy ennek a stratégiának a sikere a megfelelő antibiotikumok kiválasztásában rejlik: ha egy adott antibiotikumot annak kereszt-szenzitív partnerével váltunk fel, a kezelés sikeres lehet. Egy alternatív megközelítés lehet továbbá két olyan szer egyidejű alkalmazása, amelyek között kollaterális szenzitív kapcsolat áll fent. Ilyen módon gátolható a vad típus és a rezisztens alpopuláció növekedése és megelőzhető a rezisztencia kialakulása^{26,27}.

Források

1. Soyer, O. S. & O'Malley, M. A. Evolutionary systems biology: What it is and why it matters. *BioEssays* (2013).
2. Fehér, T., Papp, B., Pal, C. & Pósfai, G. Systematic genome reductions: theoretical and experimental approaches. *Chem Rev* **107**, 3498-3513 (2007).
3. Papp, B., Pál, C. & Hurst, L. D. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661-664 (2004).
4. Blank, L. M., Kuepfer, L. & Sauer, U. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* **6**, R49 (2005).
5. Kuepfer, L., Sauer, U. & Blank, L. M. Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome Res* **15**, 1421-1430 (2005).
6. Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., *et al.* The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362-365 (2008).
7. de Visser, J. A., Hermisson, J., Wagner, G. P., Ancel Meyers, L., *et al.* Perspective: Evolution and detection of genetic robustness. *Evolution* **57**, 1959-1972 (2003).
8. Harrison, R., Papp, B., Pál, C., Oliver, S. G. & Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci U S A* **104**, 2307-2312 (2007).
9. Hamosh, A., Scott, A. F., Amberger, J. S., Bocchini, C. A. & McKusick, V. A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic acids research* **33**, D514-D517 (2005).

10. Kondrashov, A. S., Sunyaev, S. & Kondrashov, F. A. Dobzhansky--Muller incompatibilities in protein evolution. *Proceedings of the National Academy of Sciences* **99**, 14878-14883 (2002).
11. Gao, L. & Zhang, J. Why are some human disease-associated mutations fixed in mice? *Trends in Genetics* **19**, 678-681 (2003).
12. Fisher, R. A. *The genetical theory of natural selection: a complete variorum edition* (Oxford University Press, 1930).
13. Wright, S. Surfaces of selective value revisited. *The American Naturalist* **131**, 115-123 (1988).
14. Szamecz, B., Boross, G., Kalapis, D., Kovács, K., *et al.* The genomic landscape of compensatory evolution. *PLoS biology* **12**, e1001935 (2014).
15. Esvelt, K. M. & Wang, H. H. Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology* **9**, (2013).
16. Woodruff, L. B. & Gill, R. T. Engineering genomes in multiplex. *Curr Opin Biotechnol* **22**, 576-583 (2011).
17. Pál, C., Papp, B. & Pósfai, G. The dawn of evolutionary genome engineering. *Nat Rev Genet* **15**, 504-512 (2014).
18. Carr, P. A. & Church, G. M. Genome engineering. *Nature biotechnology* **27**, 1151-1162 (2009).
19. Nyerges, Csörgő, B., Nagy, I., Latinovics, D., *et al.* Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic acids research* gku105 (2014).
20. Nyerges, Á., Csörgő, B., Nagy, I., Bálint, B., *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A* **113**, 2502-2507 (2016).
21. Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).

22. SZYBALSKI, W. & BRYSON, V. Genetic studies on microbial cross resistance to toxic agents. I. Cross resistance of *Escherichia coli* to fifteen antibiotics. *J Bacteriol* **64**, 489-499 (1952).
23. Lázár, V., Nagy, I., Spohn, R., Csörgő, B., *et al.* Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nature communications* **5**, (2014).
24. Lázár, V., Pal Singh, G., Spohn, R., Nagy, I., *et al.* Bacterial evolution of antibiotic hypersensitivity. *Molecular systems biology* **9**, (2013).
25. Dragosits, M., Mozhayskiy, V., Quinones-Soto, S., Park, J. & Tagkopoulos, I. Evolutionary potential, cross-stress behavior and the genetic basis of acquired stress resistance in *Escherichia coli*. *Molecular Systems Biology* **9**, 643 (2013).
26. Munck, C., Gumpert, H. K., Wallin, A. I. N., Wang, H. H. & Sommer, M. O. Prediction of resistance development against drug combinations by collateral responses to component drugs. *Science translational medicine* **6**, 262ra156-262ra156 (2014).
27. Baym, M., Stone, K. & Kishony, R. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science* **351**, aad3292 (2016).

Evolution and systems biology

MTA Doctoral Thesis

Csaba Pál

Synthetic and Systems Biology Unit

Biological Research Center, Hungarian Academy of Sciences, Szeged

Szeged, 2017

Integration of Mendelian genetics into evolutionary biology in the early 20th century has allowed understanding a broad range of biological phenomena, and unified several previously isolated fields including biogeography, taxonomy, ecology and population genetics. In spite of the enormous success of the Modern Synthesis, certain key issues have remained unanswered. Systems biology offers a new angle to study these problems in a consistent manner¹. In a nutshell, it aims to integrate detailed molecular network analyses and computational models of cellular subsystems with population genetics and experimental evolution, to re-investigate central issues of evolutionary biology¹.

Here I will focus on four main topics, all of which illustrate the conceptual and methodological links between evolution and systems biology. The first relates to the gene knock-out paradox. Why is it that most genes appear to be dispensable? Second, these considerations will lead to the problem of compensatory evolution. Third, I describe recent advances in bacterial genome engineering, and how genome engineering can be used to study the evolution of bacterial genomic streamlining in the laboratory. Finally, I briefly summarize our work on antibiotic resistance and collateral sensitivity in microbes.

I. Evolution of gene dispensability

In most organisms, inactivation of a single gene generally has no major effect on survival². Only 20% of the single knock-outs in yeast (*Saccharomyces cerevisiae*) are essential for growth, and similarly low figures have been observed in the worm *Caenorhabditis elegans*, *Bacillus subtilis*, and many other organisms (Table 1).

organism	no. of protein coding genes	est % of essential genes
<i>M. genitalium</i>	482	79
<i>H. pylori</i>	1590	17
<i>H. influenzae</i> Rd	1850	38
<i>S. aureus</i> N315	2594	25
<i>M. tuberculosis</i> H37Rv	4402	15
<i>B. subtilis</i>	4099	6.6
<i>E. coli</i> K12	4296	7.1
<i>S. typhimurium</i> LT2	4597	11
<i>C. glutamicum</i> R	3052	22.6
<i>F. novicida</i> U112	1719	23
<i>S. cerevisiae</i>	5794	19
<i>C. elegans</i>	19 099	7

Table 1. Distribution of essential genes in model organisms. Adapted from ref ².

Details and references therein.

These patterns tempt many issues: Is large fraction of the genes truly dispensable to the organism? Why is it that a knockout strain can grow well?

Back in 2004³, we addressed this issue using the genome scale metabolic network model of baker's yeast (*Saccharomyces cerevisiae*). The network was reconstructed from a large set of prior biochemical studies, and includes 809 metabolites connected by 851 different biochemical reactions³. Using this network, we defined a solution where fluxes of all metabolic reactions in the network satisfy the relevant constraints, given the nutrients available in the environment. Next, we calculated the optimal use of the metabolic network to produce major biosynthetic components for growth under a set of 282 predefined and ecologically relevant nutrient conditions. The model indicates that most genes have severe fitness defects

only under a small fraction (10%) of the 282 different growth conditions investigated. Thus, most genes appear to be important in specific environments only³.

Several empirical studies supported this claim. First, direct measurement of enzymatic fluxes in yeast demonstrated that about 50 percent of apparently dispensable genes are inactive under laboratory conditions^{4,5}. Even more importantly, a recent high-throughput chemogenomic study indicates that as high as 97% of the 5000 apparently nonessential genes in yeast make contribution to fitness under at least one condition⁶. Moreover, deleterious phenotypes are generally restricted to a small fraction of the tested environments⁶.

The above figures do not exclude the possibility for other mechanisms of gene dispensability⁷. Gene deletions may be compensated for by a gene duplicate with a redundant function, and reorganization of metabolic fluxes across alternative pathways may buffer gene loss^{3,7}. To approach which of the two mechanisms – gene duplicates with redundant functions versus alternative pathways – are more important, we again turned to yeast metabolism³. We focused on essential enzymatic reactions, i.e. the ones predicted to stop growth when deleted. Overall, we estimate that duplicates account for between 15–28 percent of incidences of gene dispensability, while alternative metabolic pathways can only explain 4 to 17 percent of it. These figures were later confirmed by experimental enzymatic flux measurement in the same species. These experiments suggest that, for 207 viable mutants of active reactions, network redundancy through duplicate genes is the major (75%), and alternative pathways the minor (25%) molecular mechanism of genetic network robustness. These results do not exclude the possibility that distributed robustness via alternative pathways is more common in other cellular systems.

Next we asked whether the spread and retention of a duplicate (or alternative pathway) was selected because it provided backup against mutations³. Prior theoretical works demonstrated that, under realistic mutation rates and population

size settings, most organisms are unlikely to evolve backup against mutations. So, why are duplicates in the genome? Flux balance analysis of the yeast metabolic network has shown that essential reactions are not more likely than nonessential reactions to be catalyzed by isoenzymes. Instead, isozymes appear at positions in the network where a high flux is needed. This suggests that duplicates were retained to permit a selectively advantageous increase in flux rates, a secondary consequence of which can be buffering.

Another way to ask about the evolution of distributed robustness in networks is to ask about the evolution of gene pairs that are not sequence related, but can compensate null mutations in each other. At least 51 percent of such synthetic lethal interactions are restricted to particular environmental conditions⁸. These results are compatible with a side effect model, where the enzymes are essential under nutrient specialist conditions, not because they provide backup.

II. Compensatory evolution

Genetic disorders in human populations are surprisingly frequent⁹. However, individuals carrying the same deleterious mutations often have different or no symptoms at all. Moreover, mutations deleterious in human are frequently fixed in other closely related species^{10,11}. Why is it so? I argue that evolutionary adaptation is inherently linked to the incorporation of mutations with pleiotropic side consequences. Therefore, organisms undergo major changes during evolution not simply to adapt to novel environments but also to compensate for the deleterious side-effects of adaptive mutations. Indeed, in the early 1920s Ronald Fisher pioneered the view that adaptation is by and large a hill climbing process¹²: it proceeds through progressive accumulation of beneficial mutations. By contrast, Sewall Wright proposed that fixation of conditionally deleterious mutations may act as stepping stones in evolutionary adaptation by providing access to evolutionary

pathways which are otherwise inaccessible¹³. After many decades, theoretical works on the subject are overwhelming, but the data (especially on a genomic scale) is scarce.

In our work¹⁴, we focused on a special, largely neglected aspect of this problem and asked whether deleterious gene loss events promote adaptive genetic changes and what the side consequences of such process might be¹⁴. To achieve such an ambitious goal, we integrated approaches of several disciplines, including laboratory experimental evolution and genomic analyses, coupled with bioinformatics and detailed molecular studies¹⁴. The analysis reached several important results:

Compensatory evolution following gene loss is pervasive. At least 68% of the deleterious but non-lethal null mutations can be buffered through accumulation of adaptive mutations elsewhere in the genome.

Full restoration of the lost molecular function is rare. Integrated genomic and phenotypic analyses revealed that the evolved lines diverge from each other and reach new adaptive peaks: the wild-type physiological state is generally not restored and pleiotropic side effects are prevalent.

Based on these results, we proposed that a substantial fraction of the gene content variation across species is due to the action of compensatory evolution and may not need to reflect changes in environmental conditions and consequent passive loss of genes.

III. Evolutionary genome engineering

Genome-scale engineering enables editing specific genomic locations in a directed and combinatorial manner¹⁵. Recent advances in this field offer an unprecedented opportunity to design complex molecular circuits with predefined functions¹⁶. Most

studies have either focused on engineering novel pathways that produce specific molecules for medicine and industry or attempted to construct genomic chassis that are more amenable for further rational design. We recently argued that genome engineering offers extremely powerful discovery tools for understanding the evolution of natural cellular systems¹⁷. While genome engineering had limited impact on evolutionary research so far, I predict that it will change in the near future. Technical advancements in genome engineering have the potential to transform evolutionary biology into a more predictive discipline.

Why is it important? Laboratory evolutionary experiment on microbes coupled with whole-genome sequence analyses offer powerful tools to investigate evolution in real time¹⁷. Current works largely focus on complex phenotypes of whole organisms, where genetic basis is not understood properly. However, this approach has several limitations:

- 1) **Natural genetic variation is limited in the laboratory.** Several crucial evolutionary innovations lack within population variation, on which selection could act.
- 2) **Evolution in the laboratory is slow.** Given the limited timescale of microbial laboratory evolution experiments, only relatively few mutations are fixed in most laboratory evolved populations. Therefore, comparison of these results to macroevolutionary trends is often difficult.
- 3) **No appropriate control of mutational processes.** Studying the evolution of a particular cellular subsystem is hindered by the fact that beneficial mutations can occur outside the subsystem under investigation. In other cases researchers may wish to investigate specific mutational processes only.

Genome-scale engineering provides a novel approach to study evolution in real time, as it can potentially handle the above mentioned problems¹⁸. Among others, genome engineering offers a) rapid editing and directed evolution of large genomic segments

or entire chromosomes, b) synthesis and combinatorial shuffling of small DNA segments (promoters, coding regions) or complete genomes, c) chemical synthesis and integration of large segments or even whole genomes into new host organism. For details, see ¹⁷.

Recently, we addressed some of the most long-standing problems in genome engineering^{19,20}. Currently available tools for bacterial genome manipulation suffer from three major limitations. They i) have been optimized for a few laboratory model strains (such as *Escherichia coli* MG1655), ii) demand extensive modification of the host genome prior to large-scale genome engineering, and iii) lead to the accumulation of numerous unwanted, off-target modifications, sometimes outnumbering the desired ones. Clearly, these issues have serious implications on wide-spread biotechnological applicability.

Building on prior development of multiplex automated genome engineering ²¹, our work addressed these problems and presented a simple, all-in-one solution. Briefly, we first characterized a dominant mutation in a key protein of the methyl-directed mismatch repair (MMR) system, and used it to precisely disrupt mismatch-repair in target cells²⁰. With the integration of this advance, we developed a new workflow (portMAGE) for genome-scale engineering and demonstrated its applicability for high-throughput genome editing by efficient modification of multiple loci.

Whole genome sequencing revealed that none of the modified strains carried any observable off-target mutation, a major advance over prior approaches²⁰. Due to the highly conserved nature of the bacterial MMR system, the application of dominant mutations in this system provides a unique solution to portability. By placing the entire synthetic operon that enables efficient genome engineering into a broad-host vector, we successfully adapted multiplex automated genome

engineering to a wide range of hosts, and applied the strategy for genome editing in biotechnologically and clinically relevant enterobacteria²⁰.

To demonstrate the usefulness of our system, we applied our method (pORTMAGE) to study a set of antibiotic resistance conferring mutations in *Salmonella enterica* and *E. coli*. Despite over 100 million years of divergence between the two species, mutational effects remained generally conserved, a result with implications for future systematic studies²⁰.

In sum, with just one transformation, our system allows any strain of interest across a range of enterobacteria to become an efficient host for genome-scale editing. The method simultaneously eliminates off-target mutagenesis.

Our findings have broad implications with regards to chassis engineering for the production of valuable biomaterials through the rapid optimization of biosynthetic pathways across a wide range of bacteria, a process previously requiring tedious laboratory optimization. Moreover, based on our proof-of-principle experiments, we predict that our work will open a new avenue of research in diverse fields such as functional genomics and evolutionary biology. For the first time, our method, pORTMAGE allows systematic comparison of mutational effects and epistasis across a wide range of bacterial species.

IV. Antibiotic resistance and collateral sensitivity in bacteria

Understanding how evolution of microbial resistance towards a given antibiotic enhance (cross-resistance) or decrease (collateral sensitivity) fitness in the presence of other drugs is a challenge of profound importance for several fields of basic and applied research²². Despite its obvious clinical importance, our knowledge is still limited, not least because this problem has been addressed largely by small-scale clinical studies. By combining laboratory evolution, genome sequencing, and functional analyses, our works charted the maps of cross-resistance/collateral sensitivity interactions between antibiotics in *E. coli*^{23,24}, and explored the mechanisms driving these evolutionary patterns²⁴. In a nutshell, we initiated laboratory evolutionary experiments starting with a single clone of *E. coli* K12. Parallel evolving bacterial populations were exposed to gradually increasing concentrations of one of 12 clinically relevant antibiotics, leading to up to 300-fold increase in the minimum inhibitory concentrations (MICs) relative to the wild-type. In all cases, the resistance levels were equal to or above the EUCAST clinical breakpoints. Moreover, 52% of the evolved strains showed resistance to multiple antibiotics. As a next step, the corresponding changes in susceptibilities of the lab-evolved populations were measured against a panel of other antibiotics, allowing researchers to infer a network of cross-resistance interactions. Laboratory-evolved lines were subjected to whole-genome sequence analysis and biochemical assays to decipher the underlying molecular mechanisms of these interactions.

These studies revealed that:

- a) The cross-resistance network is dense, indicating that exposure to a single antibiotic frequently yields multidrug resistance.
- b) The populations frequently evolved asymmetric cross protection, where stress A protects against stress B, but not vice versa.
- c) The network of cross-resistance is highly predictable based on antibiotic properties.

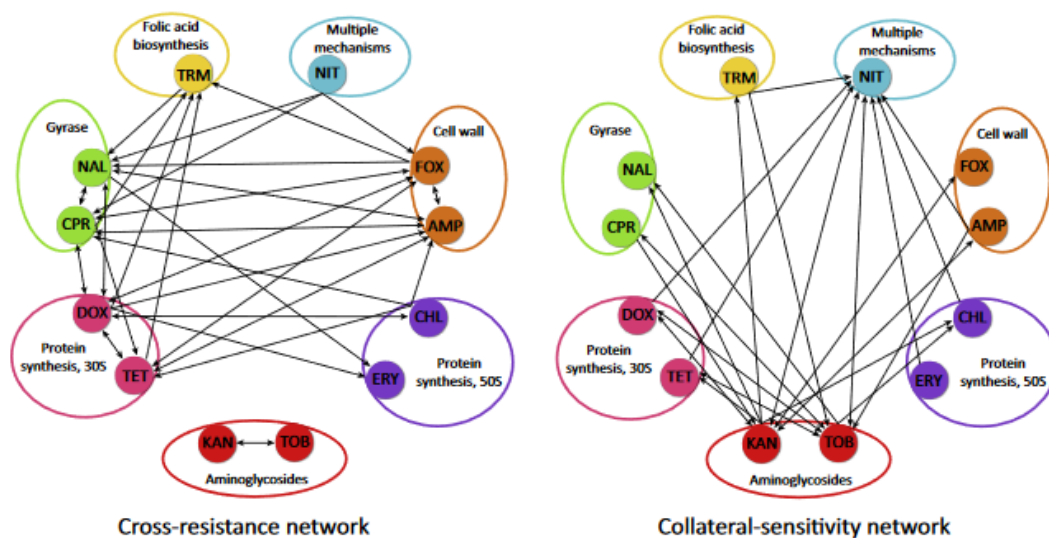


Figure 1. Based on the high-throughput measurement of antibiotic susceptibilities in laboratory evolved bacteria, two networks can be deciphered. An arrow from antibiotic A to B indicates that evolution of resistance to A generally increases (collateral sensitivity) or decreases (cross-resistance) susceptibility to B. Adapted from Pal et al. 2015.

These works also identified a strong signature of parallel evolution at the molecular level that emerged across populations adapted to different antibiotics, and such parallel mutations delivered resistance to multiple antimicrobial agents²³. The molecular mechanisms underlying antibiotic cross-resistance appeared to be very diverse, including mutations in multi drug efflux pump, metabolic genes, and genes involved in bacterial defense against a) oxidative, b) nutritional and c) membrane stresses. These works also suggested that genome-wide transcriptional rewiring mediated by global transcriptional regulatory genes has an important contribution to the cross-resistance patterns.

Perhaps the most remarkable aspect of these findings is that cross-resistance is

delivered by mutations with wide pleiotropic effects^{23,24}. Therefore, cross-protection may be more general²⁵, and opens the possibility that stressful conditions unrelated to antibiotic pressure may, as a byproduct, select for enhanced antibiotic tolerance in nature. It was indeed so, see below.

Collateral sensitivity of multidrug resistant bacteria

Prior studies demonstrated that evolution of resistance to a single antibiotic is frequently accompanied by increased resistance to multiple other antimicrobial agents. However, very little is known about the occurrence of collateral sensitivity (i.e. when evolution of resistance yields enhanced sensitivity to other antibiotics). Our studies showed that evolution of resistance towards a single antibiotic frequently yields collateral sensitivity to others²⁴. Understanding the mechanisms underlying collateral sensitivity interactions is still at an embryonic stage. We mention one example here: resistance mechanisms to various antibiotics via alteration of membrane potential have been reported in both laboratory studies and clinical settings, and such changes underlie the hypersensitivity of bacteria to other antibiotics²⁴. These results indicate the existence of antagonistic mechanisms by which bacteria modulate intracellular antibiotic concentration through altering membrane polarity²⁴.

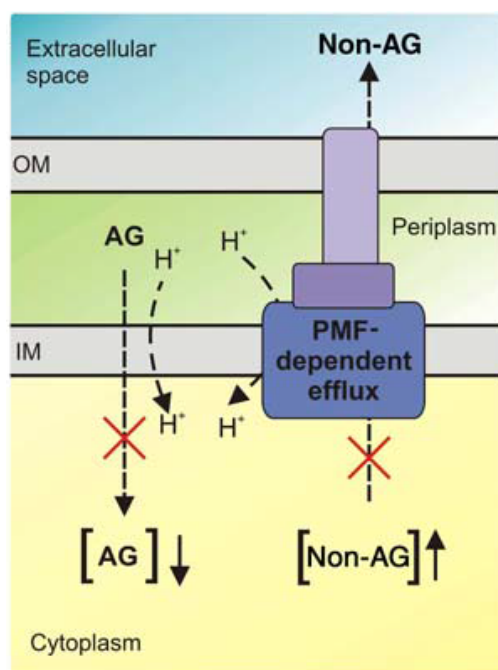


Figure 2. A mechanism underlying collateral sensitivity. Altering the membrane potential across the inner bacterial membrane has two opposing effects: it reduces the uptake of many aminoglycoside-related antibiotics but simultaneously leads to the reduced activity of PMF-dependent efflux pumps. Adapted from Lazar et al. 2013.

Clinical implications

The experimental map of cross-resistance/collateral sensitivity could serve as a unique resource and potentially permit informed decisions in medicine. For example, the choice of optimal antibiotic combinations depends on both the presence of physiological drug interactions and the availability of mutations that deliver resistance to both drugs simultaneously. It has been shown that cross-resistance between two antibiotics is largely independent of whether they show synergistic effects in combination²⁶. Combination of large-scale information on antibiotic synergism and cross-resistance could be especially informative for future development of multidrug therapies. For example, it remains controversial whether temporal rotation of antibiotics could select against the development of resistance²⁷. These works

strongly indicate that the success of such a strategy depends on the choice of antibiotics: treatment with a single antibiotic and then switching to a cross-sensitive partner may be a viable strategy. An alternative approach relies on the simultaneous administration of two agents in collateral sensitivity interaction to inhibit both the wild-type and the resistance subpopulations and thereby prevent the emergence of resistance^{26,27}.

References

1. Soyer, O. S. & O'Malley, M. A. Evolutionary systems biology: What it is and why it matters. *BioEssays* (2013).
2. Fehér, T., Papp, B., Pal, C. & Pósfai, G. Systematic genome reductions: theoretical and experimental approaches. *Chem Rev* **107**, 3498-3513 (2007).
3. Papp, B., Pál, C. & Hurst, L. D. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661-664 (2004).
4. Blank, L. M., Kuepfer, L. & Sauer, U. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* **6**, R49 (2005).
5. Kuepfer, L., Sauer, U. & Blank, L. M. Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome Res* **15**, 1421-1430 (2005).
6. Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., *et al.* The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362-365 (2008).
7. de Visser, J. A., Hermisson, J., Wagner, G. P., Ancel Meyers, L., *et al.* Perspective: Evolution and detection of genetic robustness. *Evolution* **57**, 1959-1972 (2003).
8. Harrison, R., Papp, B., Pál, C., Oliver, S. G. & Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci U S A* **104**, 2307-2312 (2007).
9. Hamosh, A., Scott, A. F., Amberger, J. S., Bocchini, C. A. & McKusick, V. A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic acids research* **33**, D514-D517 (2005).
10. Kondrashov, A. S., Sunyaev, S. & Kondrashov, F. A. Dobzhansky--Muller incompatibilities in protein evolution. *Proceedings of the National Academy of Sciences* **99**, 14878-14883 (2002).
11. Gao, L. & Zhang, J. Why are some human disease-associated mutations fixed in mice? *Trends in Genetics* **19**, 678-681 (2003).
12. Fisher, R. A. *The genetical theory of natural selection: a complete variorum edition* (Oxford University Press, 1930).

13. Wright, S. Surfaces of selective value revisited. *The American Naturalist* **131**, 115-123 (1988).
14. Szamecz, B., Boross, G., Kalapis, D., Kovács, K., *et al.* The genomic landscape of compensatory evolution. *PLoS biology* **12**, e1001935 (2014).
15. Esvelt, K. M. & Wang, H. H. Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology* **9**, (2013).
16. Woodruff, L. B. & Gill, R. T. Engineering genomes in multiplex. *Curr Opin Biotechnol* **22**, 576-583 (2011).
17. Pál, C., Papp, B. & Pósfai, G. The dawn of evolutionary genome engineering. *Nat Rev Genet* **15**, 504-512 (2014).
18. Carr, P. A. & Church, G. M. Genome engineering. *Nature biotechnology* **27**, 1151-1162 (2009).
19. Nyerges, Csörgő, B., Nagy, I., Latinovics, D., *et al.* Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic acids research* gku105 (2014).
20. Nyerges, Á., Csörgő, B., Nagy, I., Bálint, B., *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A* **113**, 2502-2507 (2016).
21. Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
22. SZYBALSKI, W. & BRYSON, V. Genetic studies on microbial cross resistance to toxic agents. I. Cross resistance of *Escherichia coli* to fifteen antibiotics. *J Bacteriol* **64**, 489-499 (1952).
23. Lázár, V., Nagy, I., Spohn, R., Csörgő, B., *et al.* Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nature communications* **5**, (2014).
24. Lázár, V., Pal Singh, G., Spohn, R., Nagy, I., *et al.* Bacterial evolution of antibiotic hypersensitivity. *Molecular systems biology* **9**, (2013).
25. Dragosits, M., Mozhayskiy, V., Quinones-Soto, S., Park, J. & Tagkopoulos, I. Evolutionary potential, cross-stress behavior and the genetic basis of acquired stress resistance in *Escherichia coli*. *Molecular Systems Biology* **9**, 643 (2013).

26. Munck, C., Gumpert, H. K., Wallin, A. I. N., Wang, H. H. & Sommer, M. O. Prediction of resistance development against drug combinations by collateral responses to component drugs. *Science translational medicine* **6**, 262ra156-262ra156 (2014).
27. Baym, M., Stone, . K. & Kishony, R. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science* **351**, aad3292 (2016).